T1-0022

AM

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/04, G01N 33/48

A1

(11) International Publication Number: **WO 98/40395**

(43) International Publication Date: 17 September 1998 (17.09.98)

(21) International Application Number: PCT/US98/04875

(22) International Filing Date: 13 March 1998 (13.03.98)

(30) Priority Data:
60/041,095 14 March 1997 (14.03.97) US

(71) Applicant *(for all designated States except US)*: TRANSGENOMIC, INC. [US/US]; 2032 Concourse Drive, San Jose, CA 95110 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GJERDE, Douglas, T. [US/US]; 12995 Woodside Drive, Saratoga, CA 95070 (US). TAYLOR, Paul, D. [US/US]; 248 Hawthorne Avenue, Palo Alto, CA 94301 (US).

(74) Agent: WALKER, William, B.; Enterprise Law Group, Inc., Suite 280, 4400 Bohannon Drive, Menlo Park, CA 94025 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: BAND ARRAY DISPLAY OF POLYNUCLEOTIDE SEPARATIONS

(57) Abstract

A method and apparatus for representing double stranded nucleic acid fragments which have been separated by a chromatographic process as an array of bands which can be accurately quantified, optimized and stored. Using, for example, a Matched Ion Polynucleotide Chromatography (MIPC) process, an analog output from a UV detector is digitized and input to a computer. The digitized signal is converted to a linear array of bands which may be displayed on a video display terminal. The intensity and/or color of a band may correlate to the amount of double stranded nucleic acid in the respective fraction or the respective double stranded nucleic acid fragment above a user selected threshold level at a corresponding point in the digitized signal. The calculated base pair length, concentration, and retention time of each band in the array of bands may be displayed in alphanumeric form.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Band Array Display Of Polynucleotide Separations

This invention relates to a method and system for improved band array displays of double stranded nucleic acid separations. More specifically, the invention is directed toward producing linear array displays of bands representing separated double stranded nucleic acid fragments.

Mixtures of double stranded nucleic acid fragments are separated for numerous and diverse reasons ranging from forensic investigations to gene mapping. The traditional and most widely used method for separating mixtures of DNA and other double stranded nucleic acid fragments is slab gel electrophoresis (GEP). Separation of double stranded nucleic acid fragment mixtures by this classical method produces a linear array of bands, wherein each band in the array represents a separated double stranded nucleic acid component of that mixture. Since many mixtures are typically separated and analyzed simultaneously in separate lanes on the same gel slab, a parallel series of such linear arrays of bands is produced. In principle, this is a highly desirable display format because it permits the observer to readily compare many separated mixtures simultaneously. The presence or absence of any band in one linear array compared to another is easily discernible if the separations are visible on the display.

However, in practice, GEP display methodology suffers from serious deficiencies which are inherent in the method.

Bands are often curved rather than straight, their mobility and shape can change across the width of the gel and lanes and bands can mix with each other. The sources of such inaccuracies stem from the lack of

uniformity and homogeneity of the gel bed, electroendosmosis, thermal gradient and diffusion effects, as well as host of other factors. Inaccuracies of this sort are well known in the GEP art and can lead to serious distortions and inaccuracies in the display of the separation results. In addition, the band display data obtained from GEP separations is not quantitative or accurate because of the uncertainties related to the shape and integrity of the bands. True quantitation of linear band array displays produced by GEP separations cannot be achieved, even when the linear band arrays are scanned with a detector and the resulting data is integrated, because the linear band arrays are scanned only across the center of the bands. Since the detector only sees a small portion of any given band and the bands are not uniform, the results produced by the scanning method are not accurate and can even be misleading.

Linear band arrays representing components of double stranded nucleic acid mixtures which have been separated by GEP have been visualized by a variety of methods, including fluorescence, direct visualization by use of a chemical stain, by adding a dye to the gel which makes the bands visible, or by tagging the DNA with radioactive P-32 before GEP separation, followed by autoradiography. These visualization methods produce a display consisting of parallel linear arrays of bands, which is a direct hard copy representation of the gel slab itself. Separation displays produced in this manner can be distorted and inaccurate because the margins of bands so displayed often are fuzzy and diffuse, rather than being sharply defined. In the autoradiography technique, for example, the radiation emitted from the separated nucleotide components in each band is omnidirectional. This

causes the exposure area of the film in contact with the gel slab to be greater than that represented by the actual band dimensions, resulting in a broadened and fuzzy display.

Analyzing adjacent bands in such displays can be a serious problem, especially when there is a large difference in the relative concentration of double stranded nucleic acid present in each band. In such a case the stronger band can obscure the weaker band and the latter may not be visible. Since the displayed data is fixed, it cannot be enhanced, optimized or manipulated, and important information can often go unnoticed. The only way to improve a defective or inadequate separation display is to re-run the separation using a more dilute sample or weaker stain. This is extremely time consuming since gel electrophoresis separations can take up to five hours or more.

A clear need, therefore, exists for an improved and flexible band array display format for double stranded nucleic acid separations in general, and DNA and RNA separations in particular, which can be electronically optimized, quantitated, and stored.

In a first aspect the invention provides a method of representing double stranded nucleic acid fragments which have been separated by Matched Ion Polynucleotide Chromatography as an array of bands, the method comprising providing a digitized signal corresponding to the double stranded nucleic acid fragments in the fractions; and converting the digitized signal into an of array of bands. In a second aspect the invention provides an apparatus for representing double stranded nucleic acid fragments which have been

separated by Matched Ion Polynucleotide Chromatography as an array of bands, the apparatus comprising acquisition means for acquiring a digitized signal, the digitized signal corresponding to the double stranded nucleic acid fragments in the fractions, conversion means, for converting the digitized signal to an array of bands corresponding to the double stranded nucleic acid fragments in the fractions, and display means for displaying the array of bands.

In a preferred embodiment of the invention, an analog signal output from the Matched Ion Polynucleotide Chromatography separation process is analog-to-digital (A/D) converted and the digitized signal is input to a computer. In the computer, the digitized signal is converted to a linear array of bands which may be displayed on a video display terminal (VDT), printer or other output device.

The bands may be displayed as lines or rectangles of fixed width. The intensity and/or color of a band may correlate to the amount of double stranded nucleic acid in the respective fraction or the respective double stranded nucleic acid fragment above a user selected threshold level at a corresponding point in the digitized signal.

Fig. **1** is a schematic representation of an embodiment of a system for producing band array displays according to the invention.

FIG. **2** illustrates a sample signal representative of an analog output from a detector in an embodiment of a system for producing band array displays according to the invention.

FIG. **3** is an illustration of the conversion of peaks in a sample signal to a band representation according to the invention.

FIG. **4** is an illustration of a gray scale band according to the invention.

FIG. **5** is a flow chart depicting an embodiment of the conversion of peaks in a sample signal to a band representation wherein the bands are represented as rectangles.

FIG. **6** is a flow chart depicting an embodiment of the conversion of peaks in a sample signal to a band representation wherein the bands are represented as lines.

FIGs. **7**, **8** and **9** are illustrations of sample displays according to the invention.

The invention produces improved linear array displays of bands representing separations of double stranded nucleic acid mixtures, wherein each band in an array represents a component fragment of the mixture. The band array displays are produced from graphical representations of digitized data.

A linear band array is a preferred format for displaying and viewing double stranded nucleic acid, and especially DNA, separation results since this is the display format of GEP, the most widely used double stranded nucleic acid separation process. A linear band array display generated from digitized data obtained from chromatographic separations, which can be accurately quantified, optimized and stored has not, heretofore been disclosed.

Recently, mixtures of double stranded nucleic acid fragments have been separated on a non polar polymeric stationary phase as described in detail in U.S. Patent 5,585,263 to G. Bonn et. al. which is incorporated by reference in its entirety herein. A major improvement in the scope and utility of the foregoing separation methodology, called Matched Ion Polynucleotide Chromatography (MIPC) has been disclosed in copending applications (Attorney Docket Nos. TRAN1-076 and TRAN1-077) which are incorporated by reference in their entirety herein. Separation of mixtures of double stranded nucleic acid fragments by MIPC may be further enhanced by removal of metal and metal ion contaminants as described in copending application (Attorney Docket No. Tran1-031)

MIPC, as used herein, is defined as a process for separating single and double stranded nucleic acids using non-polar beads having a pore size which is effective to exclude the smallest double stranded nucleic acid being separated, wherein the process uses counter ion agents and an organic solvent to desorb the double stranded nucleic acids from the beads. MIPC separates mixtures of double stranded nucleic acid fragments, double stranded DNA and RNA, on the basis of base pair length and not on the basis of nucleic acid sequence. MIPC is a size based separation. MIPC can be automated and computer controlled. Separations of DNA fragments having 5-1500 base pairs can be effectively achieved in less than 5 minutes. Such separations produce sharp and reproducible results.

MIPC was, therefore, selected as the preferred double stranded nucleic acid separation process for use with the present invention because it

provides an ideal technology for producing an accurate linear band array display format, generated from digitized data, which can be electronically optimized and quantified. Other suitable chromatographic processes capable of separating double stranded double stranded nucleic acid fragments may also be used.

FIG. **1** is a schematic representation of an embodiment of a system **2** for producing band array displays according to a first aspect of the invention. The system includes a separation column **4** coupled to a detector **6**. The detector has an analog output which is coupled to an A/D converter **8**. The A/D converter output is coupled to a CPU **10**, e.g. a personal computer (PC). The CPU includes software (not illustrated in FIG. **1**) for converting digitized data representing separated double stranded nucleic acid fragments and outputting a graphical representation of a linear array of bands wherein each band corresponds to a separated fragment of the original mixture. It is understood that close-running fragments may not be completely separated in some separations.

The CPU **10** is coupled to a storage device (e.g. hard disk, floppy disk, etc.) **16** for storing data. The CPU **10** may be coupled to a video display terminal (VDT) **18**, a printer **20**, or other output device (not illustrated) for displaying or otherwise outputting a graphical representation of separated nucleotide fragments. The CPU **10** is also coupled to a keyboard **12** and a pointing device **14**, e.g. mouse, trackball, touch pad, etc.

The preferred separation column **4** is packed with beads suitable for MIPC as described in U.S. patent applications (Attorney Docket No. TRAN1-

076) and (Attorney Docket No. TRAN1-077). The preferred detector **6** is an ultra-violet (UV) detector operating at 260 nm, the absorbance maximum of DNA. If the fragments to be detected are tagged with fluorescent or radioactive markers, a fluorescence detector or a radioactivity detector, respectively, may be used. Any detector which is capable of detecting the fragments being separated may be used in the system **2**.

The system **2** produces bands having a desired shape which provide both qualitative and quantitative information about the separated components of the double stranded nucleic acid mixture. For example, the preferred shape is a band in the form of a line or elongated rectangle of fixed width. The length, and optionally intensity, of a band, whether a line or a rectangle, is proportional to the absorbance of mixture component represented by the band. The band length, and optionally intensity, is a true and accurate quantitation of a component's relative and absolute concentration since the entire sample represented by a band is seen by the detector. This is in contrast to scanning a gel slab, wherein the detector commonly only sees the center of a band. Since the bands produced by GEP separations are not generally uniform, the data so obtained is neither quantitative nor accurate

The operation of the system **2** to produce an array of bands representing the double stranded nucleic acid separation results is as follows. A solution of a mixture of double stranded nucleic acid fragments is separated into its component fragments in the separation column **4** as described in U.S. patent applications (Attorney Docket Nos. TRAN1-076 and TRAN1-077), wherein double stranded nucleic acid fragments are separated on the basis of

size, i.e. number of base pairs. The double stranded nucleic acid mixture may be tagged with fluorescent or radioactive markers to enhance the sensitivity of detection without altering the nature of the separation.

A mobile phase solvent capable for separating the double stranded nucleic acid mixture is flowed through the column **4**. Although the exact composition of the mobile phase varies with the nature and molecular weight range of the double stranded nucleic acid mixture being separated, the most effective mobile phases contain water, an organic solvent which is completely miscible with water and a counter ion agent. Preferred organic solvents include acetonitrile, tetrahydrofuran and C-1 to C-3 alkanols. Examples of specific mobile phases which have been used to effect double stranded nucleic acid separations, including double stranded DNA separations, are described in U.S. patent applications (Attorney Docket Nos. TRAN1-076 and TRAN1-077).

The column effluent is directed to flow past the detector **6** which is coupled to the column. The detector must be capable of detecting double stranded nucleic acids or tagged analogs thereof. A preferred detector operates using a UV source and UV sensor, and detects the presence of double stranded nucleic acid fragments by measuring the change in UV absorption of the effluent as the effluent flows past the detector. If the double stranded nucleic acid fragments in the mixture to be separated have been tagged with fluorescent or radioactive markers prior to separation, then a fluorescence or radioactivity detector, respectively, can be used.

As the mobile phase flows through the column, it carries the separated double stranded nucleic acid fragments, which elute in order of lower to higher number of base pairs, through the column and past the detector. Under normal circumstances, all of the sample entering the column exits the column and travels through the detector cell. The presence of the double stranded nucleic acids fragments is detected as they flow past the detector, which responds by generating typically an analog output signal, typically an analog varying voltage. The magnitude of the output signal is a function of the quantity and absorbance of the double stranded nucleic acid present in the mobile phase passing the detector at any given time.

The analog output signal of the detector **6** is input to the A/D converter **8** where the signal is digitized, and the digitized signal is input to the CPU **10**. The rate at which the A/D converter **8** samples and digitizes the analog output of the detector **6** depends on the flow rate of the mixture through the column **4** and the detector **6**, and may optionally be varied. In a preferred embodiment, the A/D converter **8** has a variable sampling rate which can be adjusted to optimize data collection. A generally effective sampling rate is 100 millisecond intervals. If the sampling rate is varied, this is taken into account when the digitized signal is reconstructed in the computer for processing and conversion to a band representation. The digitized signal is received by the CPU **10**, and may be stored in the storage device **16** for subsequent processing. The digitized data may be displayed and/or converted and displayed in a band form in real-time and/or off line (i.e. at a later time).

All of the data from the detector are used to display the bands, even if data are manipulated so that a function is used to display some distinct property. For example, the bands may only be shown if a predetermined signal threshold is reached. However, all of the data above this threshold is used to display the band. Absolute concentrations of the separated fragments can be determined by comparing the signal corresponding to a fragment with that of an appropriate standard. For example, standards of known base pair length and concentration may be used to normalize the detector's response to the separated fragments. Optionally, if the double stranded nucleic acid fragment mixture is tagged with a fluorescent or radioactive marker, fluorescent or radioactive standards, respectively, may be used.

FIG. **2** illustrates a sample signal **100** representative of an analog output from the detector **6**. The sample signal **100** includes three peaks **102, 104, 106** which represent the output of the detector **6** in response to a mixture containing double stranded nucleic acid fragments of three different base pair lengths. The height of the sample signal **100** represents the instantaneous amount of double stranded nucleic acid fragments flowing past the detector **6** as a function of time. The three peaks **102, 104, 106** represent the amount of double stranded nucleic acid fragments of each of the three respective different base pair lengths detected in the mixture, and the total area under each of the curves of the peaks **102, 104, 106** represents the total amount of double stranded nucleic acid fragments of the three respective different base pair lengths detected in the mixture. The signal **100** illustrated in FIG. **2** and

subsequent FIGs. is by way of illustration only, and does not necessarily represent an actual signal output from a detector **6**.

FIG. **3** illustrates the conversion of the peaks **102**, **104**, **106** in the sample signal **100** to a band representation. A reconstructed signal **100'** in FIG. **3** represents a signal reconstructed from the digitized values of the original sample signal **100** (FIG. **2**). For discussion and display purposes only, the reconstructed signal **100'** is shown rotated 90 degrees from how the sample signal **100** is illustrated in FIG. **2**. In FIG. **3**, time, and hence base pair length, increases from top to bottom, and the amplitude of the reconstructed signal **100'** increases to the right.

The graphical, e.g. band, representation of the double stranded nucleic acid fragments may take several forms. For example, in a first array **110**, each of the reconstructed peaks **102'**, **104'**, **106'** is represented by a band in the form of a corresponding line **112**, **114**, **116**, with the vertical position of the lines **112**, **114**, **116** corresponding to the maximum value of the respective reconstructed peaks **102'**, **104'**, **106'**. In a second array **120**, each of the reconstructed peaks **102'**, **104'**, **106'** is represented by a band in the form of a rectangle **122**, **124**, **126**, with the vertical position and length (vertical height on the paper) of the respective rectangles **122**, **124**, **126** corresponding to a first user-selected threshold T_1. In a third array **130**, each of the reconstructed peaks **102'**, **104'**, **106'** is represented by a band in the form of a rectangle **132**, **134**, **136**, with the vertical position and length of the respective rectangles **132**, **134**, **136** corresponding to a second user-selected threshold T_2. Since T_1 is set at a higher amplitude than T_2 on the

reconstructed signal **100'**, each of the bands **122, 124, 126** in the second band array **120** corresponding to T_1 has a shorter length than the corresponding bands **132, 134, 136** in the third band array **130** corresponding to T_2. The system of the invention provides the user with the ability to select the desired form of band display, e.g. line or rectangle, and also permits the user to vary the threshold to filter out noise or other components in the reconstructed signal **100'**.

In addition, if the line or rectangular band representation is chosen, the user may select to display a gray scale, color, or other form of display, where, for example, the intensity of the gray scale, or the color, corresponds to the amplitude of the reconstructed signal. A representative gray scale band **108** is illustrated in FIG. **4**.

As also illustrated in FIG. **4**, depending on the number of quantization levels available, a gray scale band display may tend to appear as a series of discrete horizontal sub-bands. In a preferred embodiment of the invention, a "blended" gray scale band display (representing an analog variation across the band) is formed by turning individual pixels "on" or "off". In each line of pixels forming the band, the decision to turn each pixel "on" or "off" is derived by multiplying the corresponding amplitude of the reconstructed signal **100'** times a random number. If the resulting product is greater than a certain threshold, then a pixel is turned on, i.e. displayed as a black (or color) point. Otherwise, the pixel is not so displayed. Thus, as the amplitude of the reconstructed signal **100'** increases, more pixels will be displayed along corresponding lines of the respective band.

FIG. **5** is a flow chart depicting an embodiment of the conversion of peaks in a sample signal to a band representation wherein the bands are represented as rectangles. FIG. **6** is a flow chart depicting an embodiment of the conversion of peaks in a sample signal to a band representation wherein the bands are represented as lines. Both flow charts will be described as though the bands are being formed from data which has been previously collected and stored. However, those skilled in the art will readily see that the basic procedures can be used to convert sample signal data to a band representation in real time. Those skilled in the art will also recognize that the flow charts can be drawn in a number of different forms without departing from the scope of the invention.

Prior to entering either flow chart, a user will typically select a rectangle or line representation from a menu selection. Referring first to FIG. **5**, the flow begins with an INITIALIZE process block **201** in which counters, desired signal threshold, and other variables are initialized. A sample counter is then incremented by 1 in an INCREMENT COUNTER process block **202**. The current sample indicated by the sample counter is compared to the signal threshold in a SAMPLE .GE. THRESHOLD decision block **203**. If the current sample is above the threshold, then if the sample is established as the start of a new band, as determined by a NEW BAND decision block **204**, a new band is started in a START BAND process block **205**. In the START BAND **205** process block, the graphic display output for the band is started and variables are initialized for calculating sample concentration and other relevant quantities. The flow then passes through a LAST SAMPLE decision block **209**. Since the amount of eluting solvent used is sufficient to cause all the

nucleic acid mixture to pass through the column **4** and detector **6** (FIG. **1**), the LAST SAMPLE decision block **209** will always flow through its NO exit if the sample signal is above the threshold (representing fragments then being detected by the detector).

The flow then passes back up to the INCREMENT COUNTER process block **202**. Once a band has been started, i.e. the sample signal reaches or exceeds a user specified threshold, and the sample signal remains at or above the threshold, the flow from the NEW BAND **204** decision block flows through its NO exit to an EXTEND BAND process block **206**, in which the graphic display output for the band continues and the concentration and other quantities are accumulated. The process continues until the sample signal falls below the threshold which will cause the SAMPLE .GE. THRESHOLD decision block **203** to flow through its NO exit to the IN BAND decision block **207**. When the first below threshold sample is encountered after a band has been started, the IN BAND decision block **207** will flow through its YES path to an END BAND **208** process block. In the END BAND process block **208**, the graphic display output for the band is terminated at the previous sample (which was at or above the threshold), and the concentration and other relevant quantities are saved for display or other use as selected by the user.

The process continues until the SAMPLE .GE. THRESHOLD decision block **203** flows through its NO exit, the IN BAND decision block **207** flows through its NO exit, and the LAST SAMPLE decision block **209** flows through its YES exit, indicating the last sample of the run.

In the CLOSE process block **210**, data summarizing the total concentration and other relevant quantities are calculated, stored, and may be displayed.

As illustrated in FIG. **4** and described above, the bands may be displayed using color, intensity, a blended gray scale using random pixels, and the like to represent the value of the digitized detector signal. This may be implemented, for example, in each of the START BAND **205**, EXTEND BAND **206** and END BAND **208** process blocks.

Referring to FIG. **6**, the flow begins with an INITIALIZE process block **301** in which variables, counters, desired signal threshold, etc. are initialized. A maximum value is set equal to the selected threshold. A sample counter is then incremented by 1 in an INCREMENT COUNTER process block **302**. The current sample is compared to the signal threshold in a SAMPLE .GE. THRESHOLD decision block **303**. If the current sample is greater than or equal to the threshold, the current sample is then compared to the maximum value in a SAMPLE .GT. MAX decision block **304**. If the current sample is greater than the maximum value, the maximum value is then set to the value of the current sample in a MAX = SAMPLE process block **305**. The flow from the MAX = SAMPLE **305** process block proceeds to a LAST SAMPLE **308** decision block. Similar to the arrangement of FIG. **5**, the YES exit to the LAST SAMPLE **308** decision block will only be taken after all the nucleic acid mixture passes through the column **4** and detector **6**. The flow then passes back up to the INCREMENT process block **302** and continues until the current sample is less than the maximum value, indicating that the signal has

reached a peak and is beginning to decline. When this occurs, the process flow out the NO exit to the SAMPLE .GE. MAX decision block **304**.

In a DRAW BAND LINE process block **306**, a band line is displayed in a position corresponding to the peak signal value. The flow passes to a MAX = THRESHOLD process block **307** in which the maximum value is reset to the user selected threshold value.

The process continues until the last sample is reached and the LAST SAMPLE decision block **308** flows through its YES exit.

In the CLOSE process block **309**, data summarizing the total concentration and other relevant quantities are calculated, stored, and may be displayed.

In a manner similar to the display of bands as rectangles, the lines may be displayed using color, intensity, a blended gray scale using random pixels, and the like to represent the peak value of the digitized detector signal. This may be implemented in the DRAW BAND LINE process block **306**.

In the above description, the band may be displayed as a line positioned according to the peak value of the signal. In instances in which the signal is at a peak value for more than one sample, the line may be drawn in a position corresponding to the initial sample, the last sample, or at a predetermined or user selected position between the first and last samples at the peak value.

FIGs. **7**, **8** and **9** illustrate sample displays that may be output on the VDT **18**, printer **20** or other device. The sample displays in FIGs. **7**, **8** and **9** are illustrated by way of example only, and do not limit the display format or content provided by the invention. Referring to FIG. **7**, in a first sample display window **30** are a series of controls including, e.g., Peak format selection **31**, Gel format selection **32**, and an OK (or execute) **33** button. Other controls which may be included may include graphics display functions well know to those skilled in the art, e.g. zoom, pan, move (horizontal and/or vertical), rotate, copy, cut, paste, etc. Controls and/or selections may also be enabled using menu driven commands, as is also well known to those skilled in the art. Also illustrated in FIG. **7** are a series of sample band displays corresponding to the Gel format selection **32** control being selected.

The CPU **10** can store data from numerous separations. As illustrated in FIG. **7**, multiple stored separations can be displayed simultaneously for comparison of component fragments of one sample to those of another sample. When many separations are displayed simultaneously, the arrays of bands are parallel to each other as shown in Fig. **7**. If more separated samples need to be compared than can fit, e.g. on the VDT **18**, additional separations can be made to appear by scrolling using conventional techniques, e.g. the pointing device **14** or arrow or page up/page down command keys on the keyboard **12**. Referring to FIGs. **3**, **4** and **7**, bands may be represented as lines or rectangles, and may include color and/or a gray scale. In bands displayed with color and/or gray scale, the color and/or intensity of each band, may be proportional to the concentration of double stranded nucleic acid in the fraction the band represents. Any and all

separations may be displayed on the VDT **18** and/or printed by the printer **20** or other output device in any chosen display format.

Referring to FIG. **8**, in a second sample display window **30'** the Peak format selection **31** control is shown selected and the sampled signals are displayed in band format. In FIG. **9**, in a third sample display window **30"** both the Peak format selection **31** and Band format selection **32** controls are selected, and the sampled signals are displayed in both peak and band formats.

Using the zoom feature mentioned above, a particular array of bands or a segment thereof can be electronically expanded, to improve the visible resolution between adjacent bands. This feature is particularly important when one band is present in greater concentration than an adjacent band and the more concentrated band obscures the less concentrated band. In gel electrophoresis separations, a close running less concentrated band may go completely unnoticed. When multiple separations are viewed simultaneously, the lane zoom feature can be applied to all the band arrays. A similar manipulation and display may be performed in the peak format if it is chosen for display.

When a separation column is calibrated with a double stranded nucleic acid standard of known base pair length and concentration, the base pair length of the separated fragments of the mixture and their absolute concentrations can be calculated by the software. Because the separation and display methodology is reproducible, accurate base pair lengths of the

components of many mixtures can be calculated without the need to recalibrate the system for each separation.

Menu driven commands may be used to generate a variety of qualitative and quantitative information in alphanumeric form, including, but not limited to, integration of separated bands, display of base pair length, absolute concentration and relative percentage of each component fraction in a separated mixture. Bands separated from one sample may also be electronically subtracted from those of another sample to simplify visual determination of the presence or absence of a particular DNA fragment(s) in any separated DNA mixture. Similar operations may be performed if the peak shape is chosen for display.

While various embodiments and features of the invention have been described, those skilled in the art will recognize that variations and additions to those features and functions can be made within the scope of the invention. The invention is therefore intended to be limited only by the scope of the appended claims.

CLAIMS

1. A method of representing double stranded nucleic acid fragments which have been separated by Matched Ion Polynucleotide Chromatography (MIPC) as an array of bands, the method comprising

a) providing a digitized signal corresponding to the double stranded nucleic acid fragments; and

b) converting the digitized signal into an of array of bands.

2. A method of Claim 1, wherein the step of providing a digitized signal comprises

i) separating the double stranded nucleic acid fragments by MIPC;

ii) detecting the separated double stranded nucleic acid fragments; and

iii) providing a digitized signal corresponding to the detected double stranded nucleic acid fragments.

3. A method of Claim 2, wherein the substep of providing a digitized signal comprises providing an analog signal corresponding to the detected double stranded nucleic acid fragments and digitizing the analog signal.

4. A method of Claim 2 wherein the chromatographic process comprises a separation based on the number of base pairs of the double stranded nucleic acid fragments.

5. A method of Claim 2, wherein the substep of detecting the separated double stranded nucleic acid fragments comprises using an ultra violet (UV) detector.

6. A method of Claim 2, wherein the substep of detecting the separated double stranded nucleic acid fragments comprises using a fluorescence detector.

7. A method of Claim 2, wherein the substep of detecting the separated double stranded nucleic acid fragments comprises using a radioactivity detector.

8. A method of Claim 1 wherein the position of a band in the array of bands correlates to the number of base pairs of double stranded nucleic acid in a respective fragment.

9. A method of Claim 1, wherein the array of bands may be displayed as an array of lines.

10. A method of Claim 1, wherein the array of bands may be displayed as an array of rectangles.

11. A method of Claim 9, wherein a band is displayed in color, and wherein the color of the band correlates to the amount of double stranded nucleic acid fragment represented by that band.

12. A method of Claim 9, wherein a band is displayed in a gray scale, and wherein the level of the gray scale at a point along the length of the band correlates to the amount of the respective double stranded nucleic

acid fragment above a threshold level at a corresponding point in the digitized signal.

13. A method of Claim 9, wherein the length of a rectangular band correlates to the presence of a level of the respective double stranded nucleic acid fragment above a user selectable threshold level.

14. A method of Claim 1, wherein the array of bands may be displayed simultaneously with its corresponding digitized signal.

15. An apparatus for representing double stranded nucleic acid fragments which have been separated by MIPC as an array of bands, the apparatus comprising

a) acquisition means for acquiring a digitized signal, the digitized signal corresponding to the double stranded nucleic acid fragments,

b) conversion means, for converting the digitized signal to an array of bands corresponding to the double stranded nucleic acid fragments, and

c) display means for displaying the array of bands.

16. An apparatus of Claim 15, wherein the acquisition means comprises an A/D converter.

17. An apparatus of Claim 16, wherein the acquisition means comprises a MIPC separation column having a double stranded nucleic acid detector in communication therewith, the double stranded nucleic acid detector having an analog output coupled to the A/D converter.

18. An apparatus of Claim 17, wherein the double stranded nucleic acid detector comprises a UV detector.

19. An apparatus of Claim 17, wherein the double stranded nucleic acid detector comprises a fluorescence detector.

20. An apparatus of Claim 17, wherein the double stranded nucleic acid detector comprises a radioactivity detector.

21. An apparatus of Claim 15, wherein the conversion means comprises a CPU for converting the digitized signal into a linear array of bands.

22. An apparatus of Claim 15, wherein the display means comprises means for the display of multiple linear arrays of bands.

23. An apparatus of Claim 15 comprising calculating means for calculating the number of base pairs, concentration and retention time of each double stranded nucleic acid fragment represented by each corresponding band the array of bands.

24. An apparatus of Claim 15 comprising alphanumeric display means for the alphanumeric display of the calculated base pair length, concentration, and retention time of each band in the array of bands.

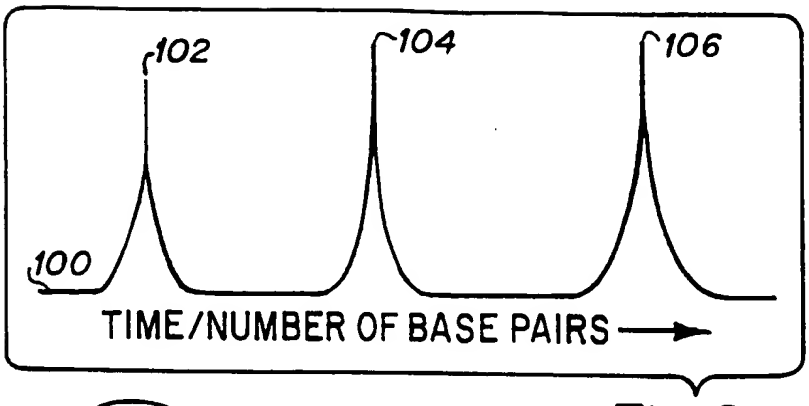

Fig_2

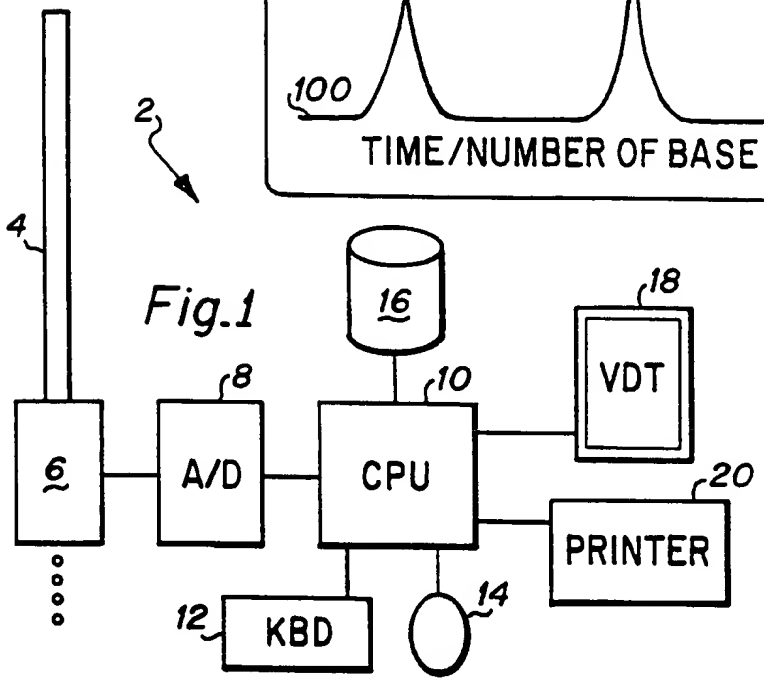

Fig_1

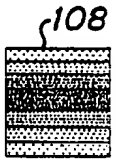

Fig_4

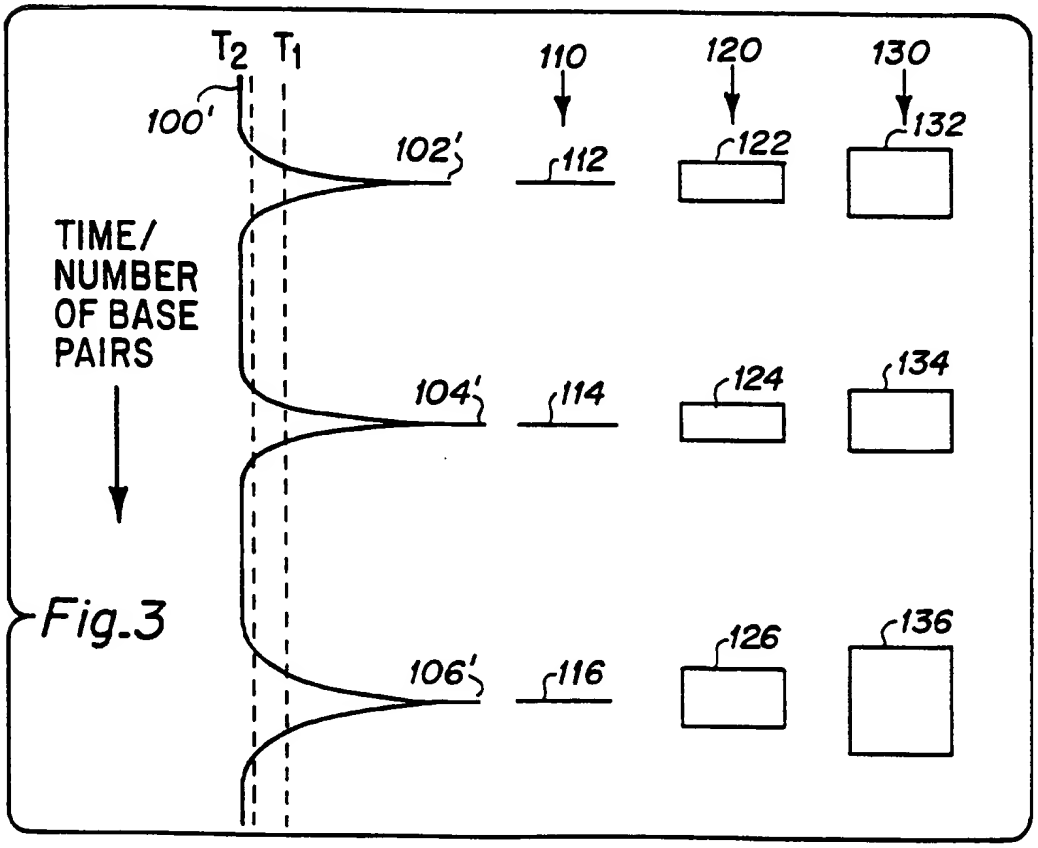

Fig_3

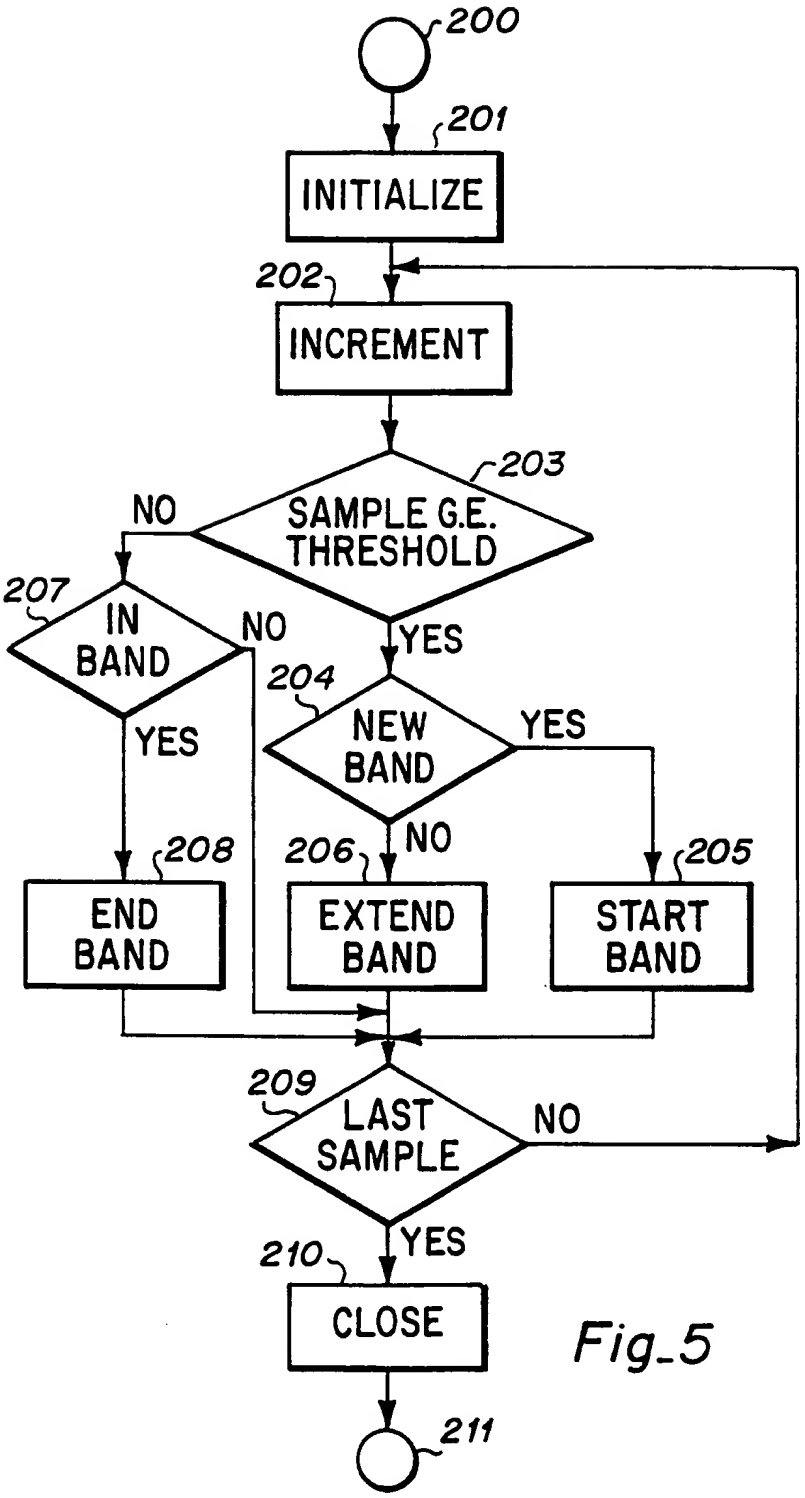
200
201
INITIALIZE
202
INCREMENT
203
SAMPLE G.E. THRESHOLD
NO
YES
207
IN BAND
NO
YES
204
NEW BAND
YES
NO
208
END BAND
206
EXTEND BAND
205
START BAND
209
LAST SAMPLE
NO
YES
210
CLOSE
211

Fig_5

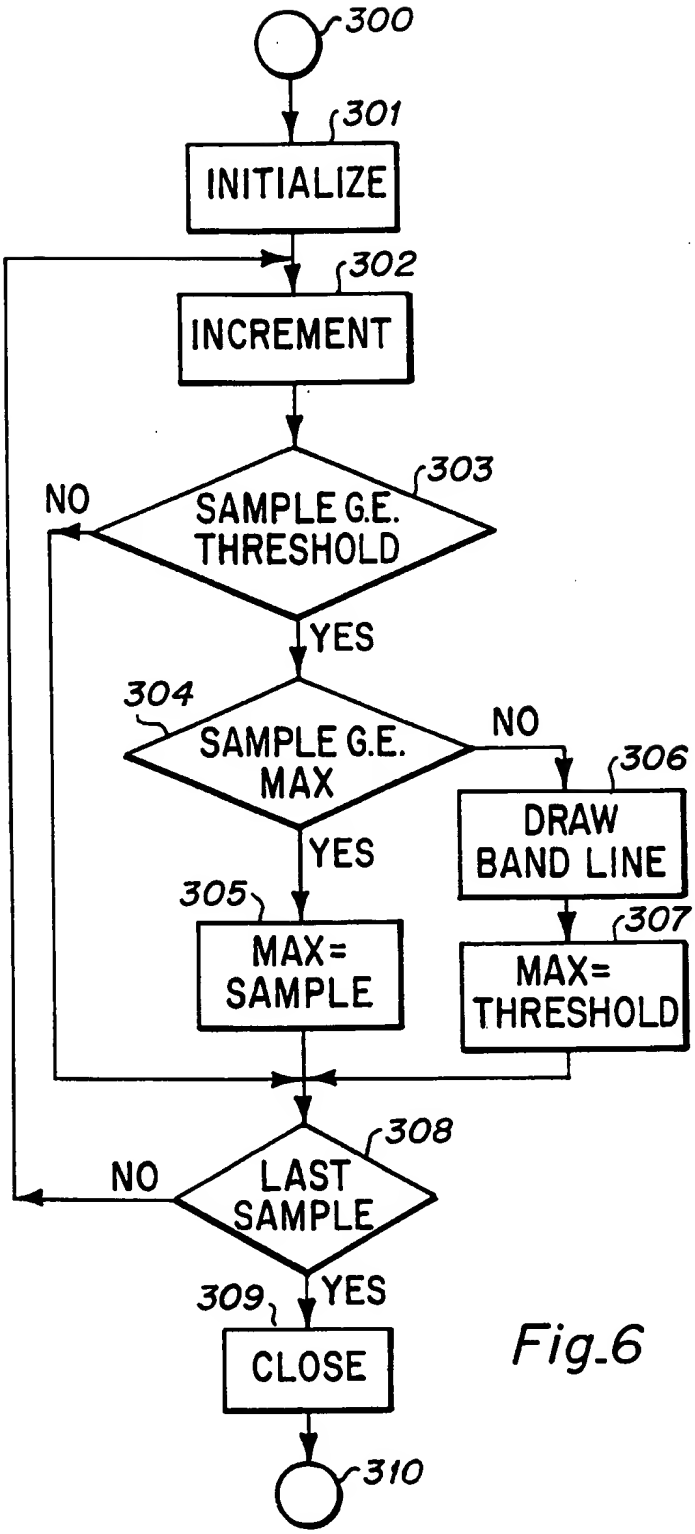
300
301
INITIALIZE
302
INCREMENT
303
SAMPLE G.E. THRESHOLD
NO
YES
304
SAMPLE G.E. MAX
NO
306
DRAW BAND LINE
YES
305
MAX= SAMPLE
307
MAX= THRESHOLD
308
LAST SAMPLE
NO
YES
309
CLOSE
310

Fig.6

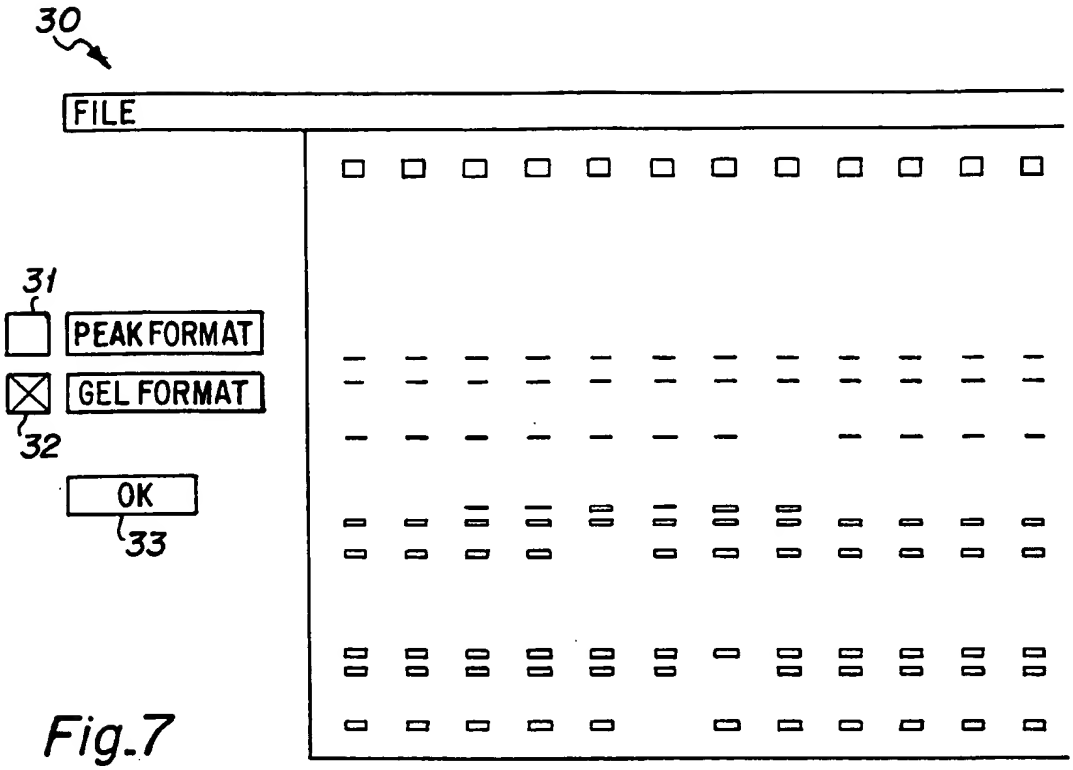
30
FILE
31
PEAK FORMAT
GEL FORMAT
32
OK
33

Fig.7

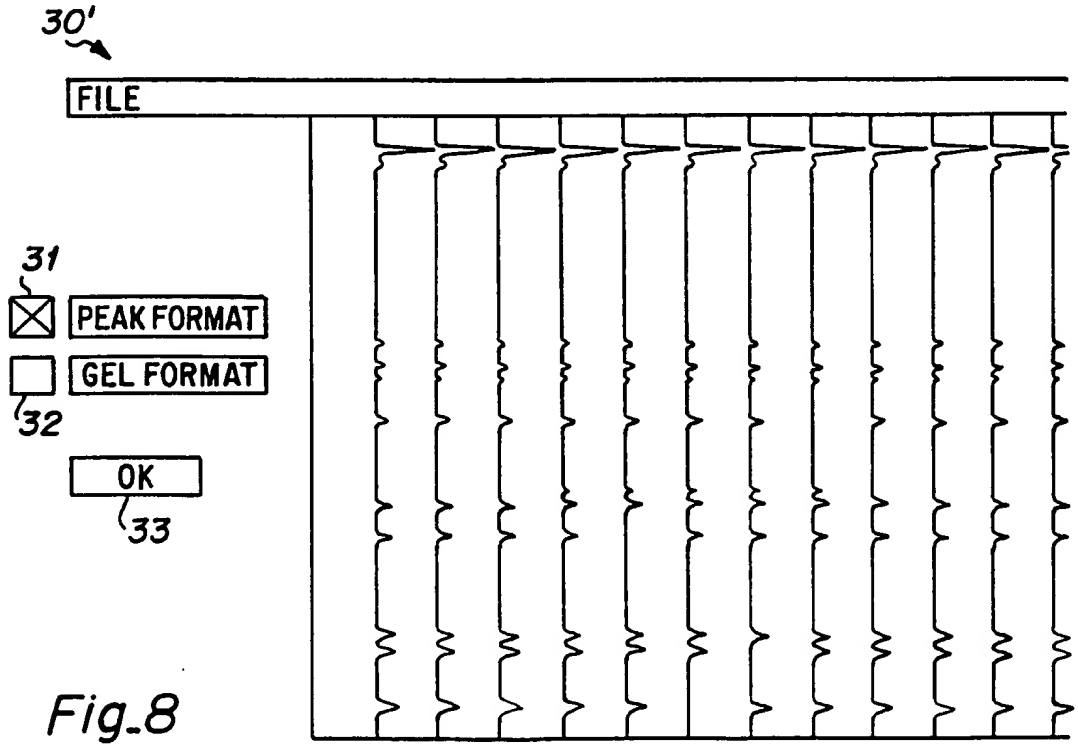
30'
FILE
31
PEAK FORMAT
GEL FORMAT
32
OK
33

Fig.8

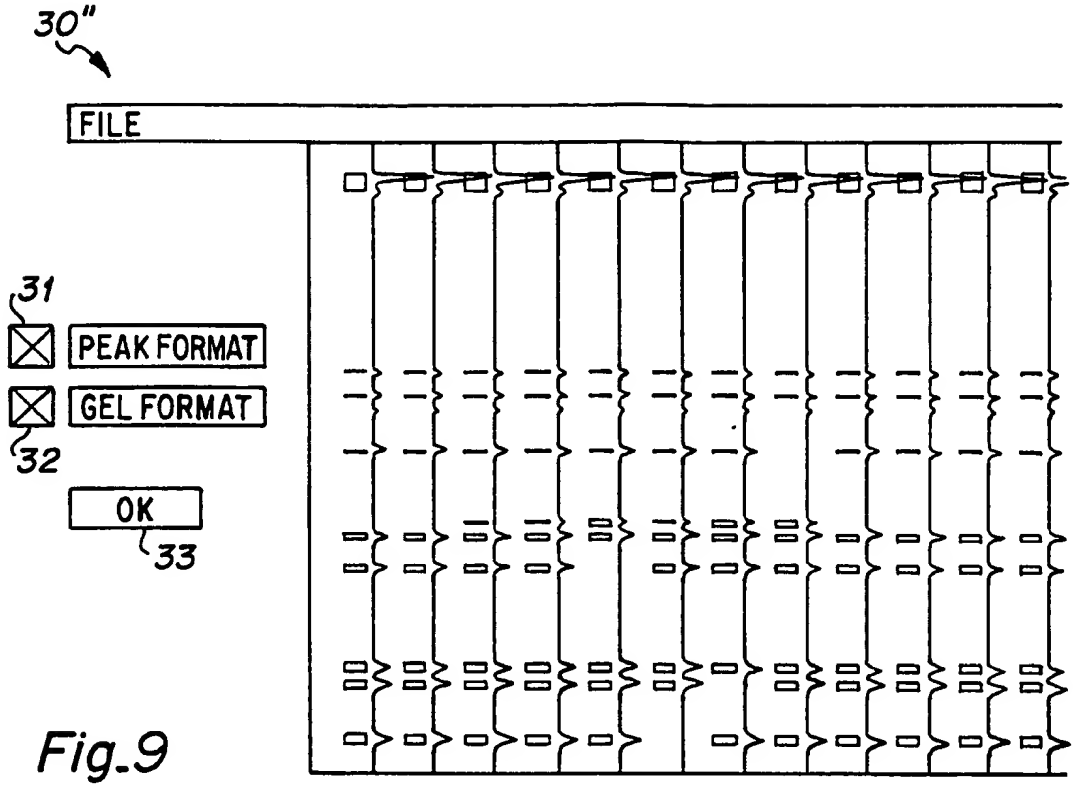

Fig.9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/04875

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 21/04; G01N 33/48

US CL :435/288.6, 288.7; 436/94; 536/25.4; 422/59, 68.1, 82.05, 101

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/288.6, 288.7; 436/94; 536/25.4; 422/59, 68.1, 82.05, 101

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, BIOSIS, MEDLINE, CA, DERWENT

search terms: chromatography, digitized, analog, converter, fragments, bands, array, dna, nucleic acid

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,585,236 A (BONN et al.) 17 December 1996, col. 2 line 65 to col. 3 line 26.	1-24
Y	US 4,927,265 A (BROWNLEE) 22 May 1990, col. 1 lines 20-30; col. 7 lines 12-28; col. 9, lines 17-27	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 22 APRIL 1998	Date of mailing of the international search report 13 JUL 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer KENNETH R. HORLICK Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*